

2

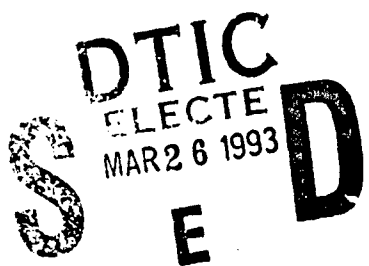
AD-A262 558



Technical Report
970

Programmable Interactive System for Cochlear Implant Electrode Stimulation

Reproduced From
Best Available Copy



J. Tierney
M.A. Zissman
D.K. Eddington
W.M. Rabinowitz

12 January 1993

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS



Prepared for the Department of the Air Force
under Contract F19628-90-C-0002.

Approved for public release; distribution is unlimited.

20000929083

98 3 25 159

93-06219



This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. The work was sponsored by the Department of the Air Force under Contract F19628-90-C-0002.

This report may be reproduced to satisfy needs of U.S. Government agencies.

The ESC Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER


Gary Tutungian
Administrative Contracting Officer
Directorate of Contracted Support Management

Non-Lincoln Recipients

PLEASE DO NOT RETURN

Permission is given to destroy this document
when it is no longer needed.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

**PROGRAMMABLE INTERACTIVE SYSTEM FOR
COCHLEAR IMPLANT ELECTRODE STIMULATION**

**J. TIERNEY
M.A. ZISSMAN
Group 24**

**D.K. EDDINGTON
W.M. RABINOWITZ
MIT Research Laboratory of Electronics
and Massachusetts Eye and Ear Infirmary**

TECHNICAL REPORT 970

12 JANUARY 1993

Approved for public release; distribution is unlimited.

LEXINGTON

MASSACHUSETTS

ABSTRACT

The aim of this research, which was performed as a Lincoln Laboratory Innovative Research Program (IRP) project, was to apply advanced digital speech and signal-processing techniques toward improving cochlear implant electrode stimulators. By providing a flexible stimulator whose function could be tuned depending on the subject's residual auditory nerves and the efficiency of the implant's coupling to those nerves, it was hypothesized that the subject's speech reception could be improved. The approach to providing these new and improved electrode stimulators included the design of a laboratory signal processor used for interactive testing of new algorithms with implant subjects. This Programmable Interactive System for Cochlear Implant Electrode Stimulation (PISCES) was designed, built, and tested at Lincoln Laboratory and then delivered to the Massachusetts Eye and Ear Infirmary (MEEI) Cochlear Implant Research Laboratory (CIRL). In collaboration with researchers at MEEI CIRL and MIT Research Laboratory of Electronics (RLE), new algorithms run on PISCES have resulted in substantial improvements in subject speech reception relative to that with their current implant stimulators. These results were obtained as a result of interactive algorithm adjustment at the clinic, which demonstrated the importance of a flexible signal processor.

DTIC QUALITY INSPECTED 1

Accession For	
NTIS CRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

ACKNOWLEDGMENTS

This Lincoln Laboratory Innovative Research Program (IRP) project required close cooperation between Lincoln Laboratory and Massachusetts Eye and Ear Infirmary personnel. The authors also wish to thank Dr. Joseph P. Donnelly, Dr. Ronald R. Parenti, and the rest of the IRP committee for their belief that this was a worthwhile project. Lincoln Division Heads Alan J. McLaughlin and Peter E. Blankenship and Group Leaders Dr. Clifford J. Weinstein and Gerald C. O'Leary were very flexible in allowing us to proceed with this project, providing Group 24 resources as required. We especially appreciate Peter's efforts to help us secure follow-up support from NIH, which will fund conversion of PISCES from a laboratory system to a subject-wearable device. Finally, we are grateful for the patience and flexibility of all our subjects.

TABLE OF CONTENTS

Abstract	iii
Acknowledgments	v
List of Illustrations	ix
List of Tables	xi
1. INTRODUCTION	1
2. THE COCHLEAR IMPLANT FOR SENSORY/NEURAL DEAFNESS	3
3. THE PISCES SYSTEM	7
3.1 Hardware Configuration	7
3.2 Software Environment	11
4. ALGORITHMS	15
4.1 Digital Simulation of the Ineraid Stimulator	15
4.2 Continuous Interleaved Sampling Algorithm	21
5. CLINICAL EXPERIENCES AND RESULTS	29
5.1 Psychophysical Measurements	30
5.2 Speech Materials	30
5.3 Subject Experiments	31
5.4 Comments on the Clinical Experience	38
6. CONCLUSIONS	39
APPENDIX A - The Advantages of Quadrature Envelope Estimation	41
REFERENCES	45

LIST OF ILLUSTRATIONS

Figure No.		Page
1	Block diagram of the peripheral auditory system.	3
2	The cochlear implant in place.	4
3	Patterns of nerve and hair cell survival for a normal and deaf ear.	5
4	Block diagram of the PISCES hardware.	7
5	Photograph of the PISCES hardware.	8
6	Block diagram of the Sonitech Spirit-30 card.	10
7	Block diagram of the three operating modes.	11
8	Signal processing performed by the Ineraid hardware stimulator and the digital simulation.	15
9	A cbank main parameter specification file.	18
10	A cbank dynamic range compression file.	18
11	Example of dynamic range compression curve.	19
12	Example of cbank band-pass filters.	20
13	cbank input and output. Top figure shows waveform for input sentence: "We finished the IRP." Middle figure shows four channels of corresponding cbank output. Bottom figure shows a wideband spectrogram of the input.	21
14	The continuous interleaved sampling stimulator.	22
15	Pulse outputs from pbank .	23
16	A pbank main parameter specification file.	25
17	pbank input and output. Top figure shows a waveform for input sentence: "Massachusetts Eye and Ear Infirmary." Next figure shows estimated envelopes for all six channels. Third figure shows a close-up of the pulse output in the middle of an unvoiced fricative. Bottom figure shows wideband spectrogram of input.	28
18	Photograph of PISCES and current isolator/limiter equipment at MEEI CIRL.	29
19	A typical output compression curve.	31
20	Filter bank approximating Butterworth filters.	33
21	Filter bank with more rectangular filters.	34

LIST OF TABLES

Table No.		Page
1	The cbank Command Line Arguments	17
2	The pbank Command Line Arguments	24
3	Dynamic Range Measurements for Subjects S04 and S05	32
4	Subject S04 Performance on 24 VCV Test	32
5	Subject S05 Performance on 16 VCV Test	35
6	Best Performance for Six Subjects	36

1. INTRODUCTION

Over 300,000 people in the United States suffer from a profound hearing loss. In these cases, treatment via conventional hearing aids is ineffective. In most of these cases, an array of electrodes can be surgically implanted to excite surviving inner ear auditory neurons. These electrodes are stimulated by devices that transduce acoustic waves (speech, music, noise, etc.) to electric signals. The signal processing performed by the electrode stimulators is more complex than the frequency dependent amplification performed by a conventional hearing aid.

The aim of this Innovative Research Program (IRP) project was to apply Lincoln Laboratory digital speech and signal-processing expertise toward improving cochlear implant electrode stimulators.¹ By providing a flexible stimulator whose function could be tuned depending on the subject's residual nerves and the efficiency of the implant's coupling to those nerves, it was hypothesized that speech reception could be improved. The approach to providing these new and improved electrode stimulators included the design of a laboratory signal processor used for interactive testing of new algorithms with implant subjects. This Programmable Interactive System for Cochlear Implant Electrode Stimulation (PISCES) was designed, built, and tested at Lincoln Laboratory and then delivered to the Massachusetts Eye and Ear Infirmary (MEEI) Cochlear Implant Research Laboratory (CIRL). In collaboration with researchers at MEEI CIRL and MIT Research Laboratory of Electronics (RLE), new algorithms were designed and run on PISCES and have resulted in speech reception improvements for implant subjects relative to their current implant stimulators. These improvements were obtained as a result of interactive algorithm adjustment at the clinic, which demonstrated the importance of a flexible signal processor.

This report summarizes the design, implementation and testing activities of the IRP project. Chapter 2 describes the cochlear implant and the conditions that make it necessary. Chapter 3 outlines the PISCES hardware and software design. Chapter 4 details the algorithms that have been implemented on PISCES and used in clinical interactions. Chapter 5 reports the results of testing with six implant subjects. Finally, Chapter 6 discusses the consequences of this IRP effort and the subsequent follow-on work.

¹In this report, the phrases "electrode stimulator" and "implant stimulator" refer to all of the processing that converts an acoustic signal to a current source output used to drive an implant electrode. The term "processor" has been avoided as it can denote a laboratory-based computer, a microcomputer (DSP chip) within the laboratory computer, a portable analog or digital acoustic-to-current transducer, or an algorithm running in a digital signal processor (either laboratory based or portable).

2. THE COCHLEAR IMPLANT FOR SENSORY/NEURAL DEAFNESS

In the healthy human peripheral auditory system, sound perception begins when an incident acoustic wave causes the ear drum to vibrate [2]. As shown in Figure 1, this vibration is coupled

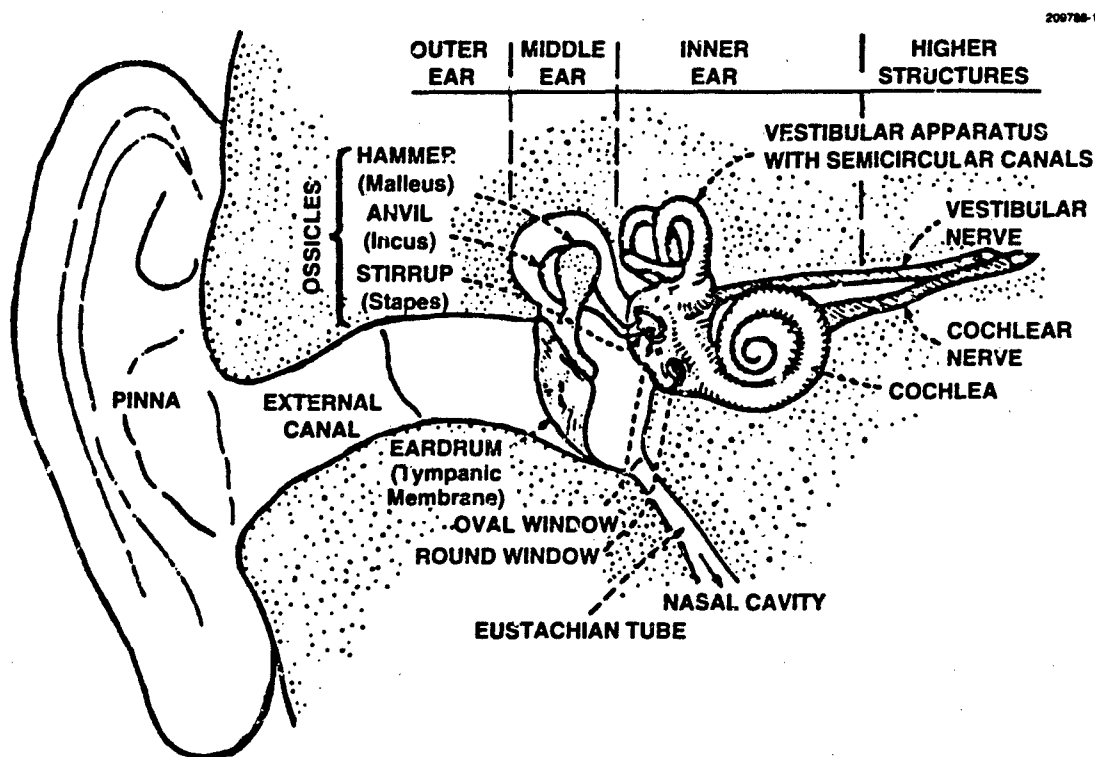


Figure 1. Block diagram of the peripheral auditory system.

through three small bones in the middle ear to the cochlea of the inner ear. The cochlea is a helical structure that is surrounded by bone and filled with fluid. The basilar membrane extends along the cochlea for about 35 mm in the same helical shape. Sound vibrations that are conducted through the middle ear to the cochlea cause the basilar membrane to vibrate. The end of the basilar membrane near the coupled input at the oval window responds to high frequencies, while the end farthest from the input responds to the lowest frequencies. Along the basilar membrane is the organ of Corti, an elaborate structure containing thousands of hair cells. The bending motion

of the basilar membrane causes small cilia on these hair cells to release neurotransmitters which excite nearby auditory neurons. Although the auditory neurons transmit a rich variety of signals to the brain, the most basic parameter is simply the place at which the neuron originates on the membrane. Sensing which neurons are active is one cue the brain uses to map the frequency content of acoustic signals. One measure of signal intensity is the total neuronal activity. Shown in Figure 2 are the details of the cochlea in cross section. The tympanic canal is the passage to which

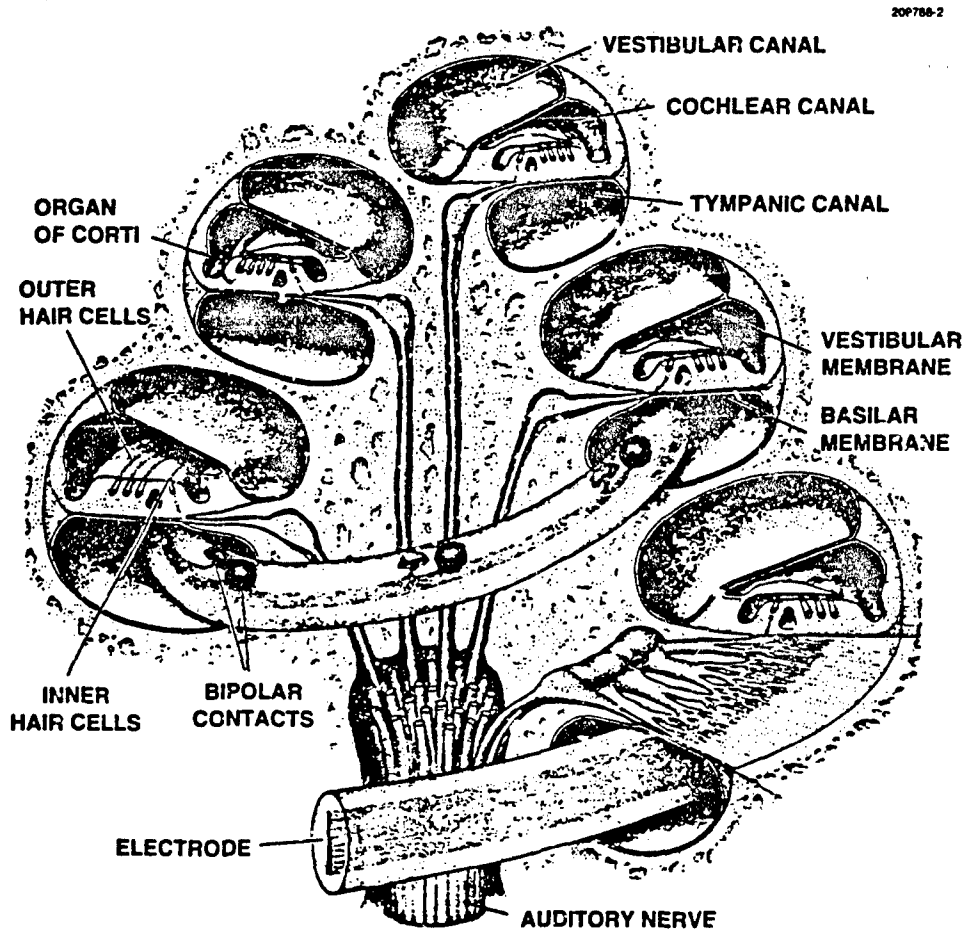


Figure 2. The cochlear implant in place.

the round window allows pressure to be released, while the vestibular canal is terminated by the oval window driven from the stapes. These two spaces are separated by a third space called the medial or cochlear canal. The frequency selective basilar membrane forms one wall of this medial space which also contains the organ of Corti as well as the auditory neurons. The 30,000 auditory neurons provide all the information used by the brain to perceive the acoustic environment.

One common form of deafness is associated with the gradual deterioration of the middle ear coupling bones [14]. Because this deterioration results in attenuation in the path from acoustic input to the basilar membrane, this condition is sometimes treated with conventional (i.e., linear filtering) hearing aids or surgery. Another source of severe hearing impairment involves the loss of sensory hair cells and/or the sensory neurons connected to them. This form of hearing loss may be caused by extremely loud sounds that damage the sensitive hair cell cilia, by the effect of certain drug treatments on the cilia, by disease which may destroy the hair cells and the neurons, or by the lack of transducer mechanisms due to a congenital condition. About 300,000 cases of sensory hearing impairment exist in the United States alone. The left-hand panel of Figure 3 shows a healthy array of hair cells and neurons. Hair cells are shown as loops on the basilar membrane, and neurons are shown connecting hair cells to the auditory nerve. The auditory nerve connects to the brain stem. In contrast, the right-hand panel of Figure 3 shows degeneration associated with sensory/neural deafness.

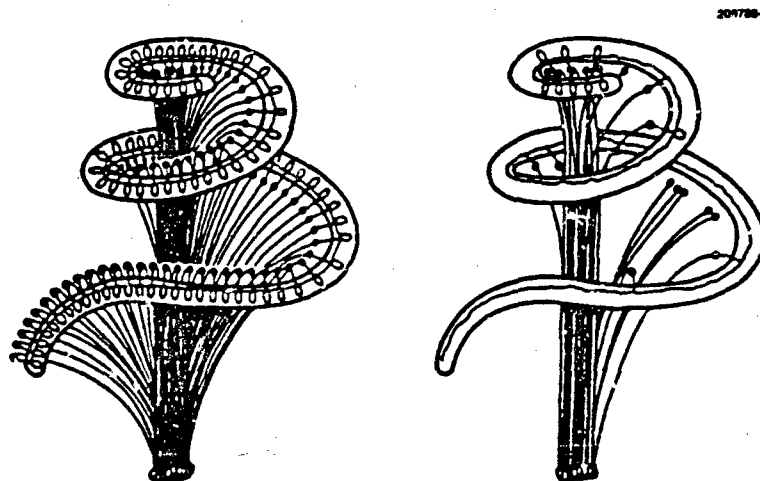


Figure 3. Patterns of nerve and hair cell survival for a normal and deaf ear.

Cochlear implants were developed to substitute for damaged hair cells and other elements in the cochlea and to bypass the external and middle ear pathways [4]. The implant is an electrode array that provides electric fields in close proximity to the remaining nerve fibers when excited by a stimulator. The electrode stimulator is a signal-processing device that converts incoming acoustic signals into stimuli appropriate for the implant electrodes. A survey of cochlear implants and stimulators was undertaken by Ifukube [5].

The insertion of a cochlear implant is a surgical procedure that varies with the type of electrode array to be used as well as the place at which it is meant to establish electrical currents and fields. Early implants comprised a single electrode in the middle ear close to the cochlea, while modern procedures snake an electrode assembly into the cochlear cavity to interact with a wide range of remaining sensory neurons [7]. Figure 2 shows such an electrode assembly inserted through the membrane of the round window. The electrode assembly comprises multiple insulated wires. Each wire is connected at one end to a conducting contact placed near the basilar membrane. These contacts are distributed along the length of the basilar membrane. The other end of each insulated wire connects eventually to the electrode stimulator. The electrodes can be excited as balanced pairs or as monopolar electrodes with a common return.

The human subjects tested as part of this project have had Ineraid [3] electrode arrays surgically implanted. These implants consist of six electrodes distributed along the first 20 to 24 mm of the cochlea and two extracochlear electrodes that can be used as ground returns. Each subject wears on his belt an Ineraid four-channel stimulator designed to excite four of the implant electrodes. During visits to CIRL, the Ineraid stimulator was replaced by PISCES, the programmable, interactive system for cochlear implant electrode stimulation described in Chapter 3.

3. THE PISCES SYSTEM

This chapter describes the PISCES hardware and software used to convert acoustic speech input into signals suitable for stimulating cochlear implant electrodes. To test new implant electrode stimulation strategies, PISCES was used in place of the Ineraid hardware stimulator in laboratory experiments.

3.1 Hardware Configuration

A block diagram of the PISCES hardware is shown in Figure 4 and a photograph is shown in

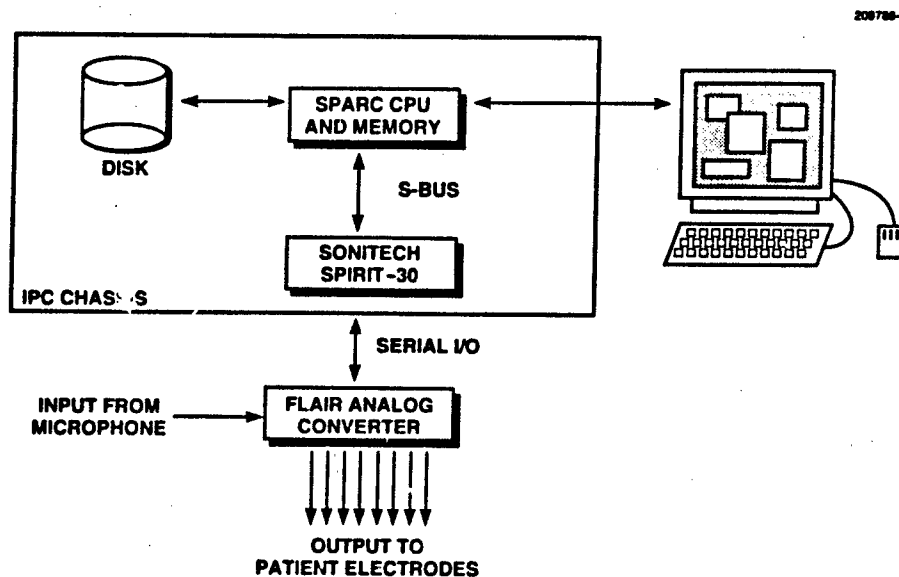


Figure 4. Block diagram of the PISCES hardware.

Figure 5. The following sections outline the major PISCES hardware modules.

3.1.1 Host Computer

The host computer fulfills many objectives. First, it functions as a general-purpose computer, performing non-real-time floating-point electrode stimulation simulations. Next, it gives the user the ability to review graphically the effects of a given stimulation algorithm both for debugging

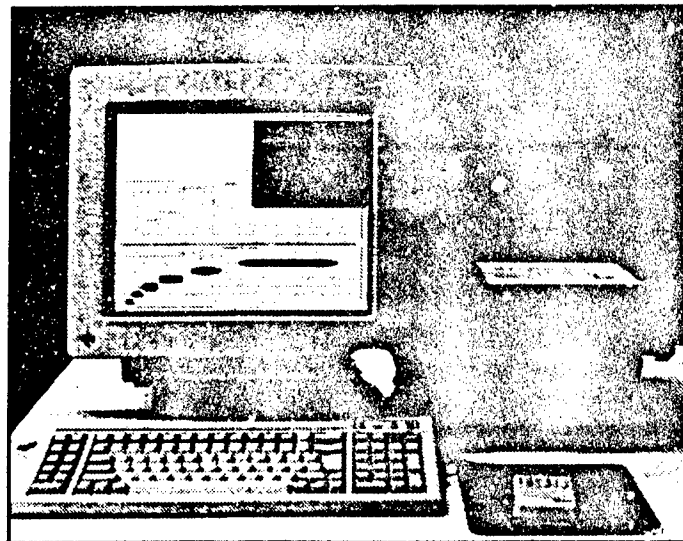


Figure 5. Photograph of the PISCES hardware.

purposes and for comparison against other algorithms. Additionally, when the host computer controls a special-purpose DSP board, its tasks include downloading the stimulation algorithms and parameters to the DSP board; initializing and interrupting board processing; and passing data among the user, the DSP board, and the disk file system.

After considering both UNIX workstations and IBM-PC/AT compatibles, a Sun Microsystems SPARCstation IPC was chosen to serve as the PISCES host. The IPC is an inexpensive, 4.2-MFLOPS UNIX workstation incorporating a 25-MHz SPARC integer and floating-point processor. It has spare S-Bus slots for installation of peripheral hardware. With regard to software, the IPC comes loaded with SunOS (Sun's version of UNIX) and OpenWindows (Sun's version of X-windows), thereby affording a multiuser, multitasking environment not commonly available on IBM-PC/AT compatibles. Although DSP processing boards are much more widely available for the IBM-PC/AT platform, the Sun/UNIX expertise accumulated by the Lincoln and MEEI personnel prior to this project heavily influenced the selection of the IPC.

3.1.2 DSP Board

Several DSP boards were commercially available for incorporation into the IPC. Each board contains a single DSP chip, fast memory, serial interfaces, and an S-Bus interface to the IPC. Boards containing the Motorola 56000, AT&T DSP32C, and Texas Instruments TMS320C30 were

considered. Boards using the 56000 were disqualified early on, as the 56000 is a fixed-point processor. Given the availability of fast, floating-point processors, there was a desire to bypass the complications of fixed-point arithmetic. The DSP32C was considered but disqualified due to the explicit pipelined nature of its assembly code. Although it was anticipated that much of the software would be written in C, it seemed inevitable that some assembly coding would be required, and it was the experience of the Lincoln personnel that the DSP32C is not easily programmed in assembly language. The TMS320C30, a floating-point chip that is easily programmed in assembly language, had been employed successfully for other projects at Lincoln; thus, it was chosen as the PISCES DSP chip.

At the onset of the IRP project, only one DSP board vendor manufactured a TMS320C30-based board compatible with the IPC. Sonitech Incorporated's Spirit-30 S-Bus card comprises a 33-MFLOPS TMS320C30, an S-Bus interface, 2 Mbytes of zero wait state RAM, and two serial ports.² Furthermore, the Spirit-30 supports the SPOX operating system (see Section 3.2 below), which eases the software migration from non-real-time workstation simulations to real-time DSP board implementations. Figure 6 shows a block diagram of the Spirit-30 card.

²Subsequently, Loughborough Sound Images has introduced an S-Bus-based C30 card quite similar to the Spirit-30.

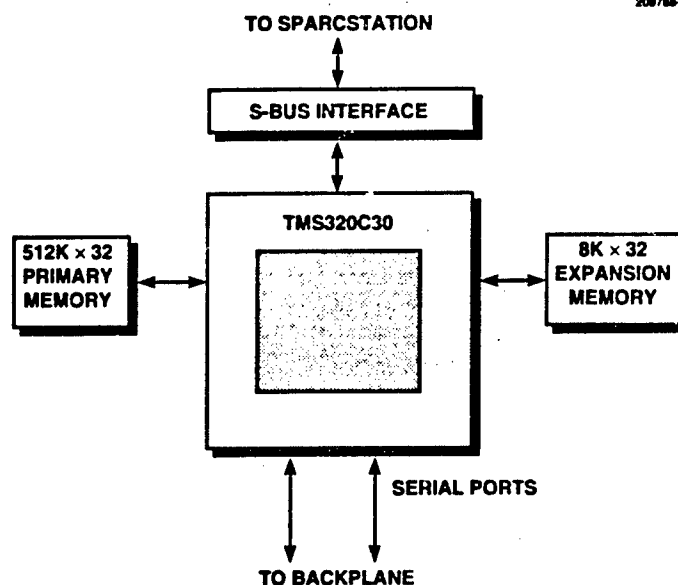


Figure 6. Block diagram of the Sonitech Spirit-90 card.

3.1.3 Analog Interface

To provide an A/D and D/A capability for the Texas Instruments TMS320C30 DSP chip used by the Sonitech board, the Flexible Lincoln Audio Interface (FLAIR) board was designed to connect to the serial I/O port of the TMS320C30 chip. The FLAIR board provides a two-channel A/D input stream using the Crystal CS5336 16-bit delta-sigma modulation converter. As the chip performs oversampling followed by digital filtering and downsampling, it exhibits a high signal-to-noise ratio over a wide range of sampling rates. This approach eliminates the need for external antialiasing filters.

The D/A output from the FLAIR board is provided by Burr Brown PCM56P 16-bit converters that have a settling time of 1.5 μ s. Between the C30 chip and the D/A converters, Nippon Precision Circuits SM5813AP upsampling FIR filter chips can be optionally engaged. Each SM5813AP provides two channels of a 1:8 upsampling and associated digital low-pass filtering. Upsampling and filtering prior to D/A conversion eliminates the need for a sharp analog smoothing filter. The converters can also be driven directly from the C30, bypassing the upsampling and filtering stages, to generate pulse signals at the sampling interval width.

The FLAIR board assembly that is used in PISCES provides two channels of 16-bit A/D input and eight channels of 16-bit D/A output. Sampling rate, use of upsampling or direct outputs, and

the number of A/D and D/A channels are specified under software control. The hardware assembly can be easily extended to provide additional D/A outputs.

3.2 Software Environment

Development of implant stimulation software on PISCES generally takes advantage of the three processing modes shown in Figure 7 and described below. Detailed descriptions of the actual stimulation algorithms are postponed until Chapter 4.

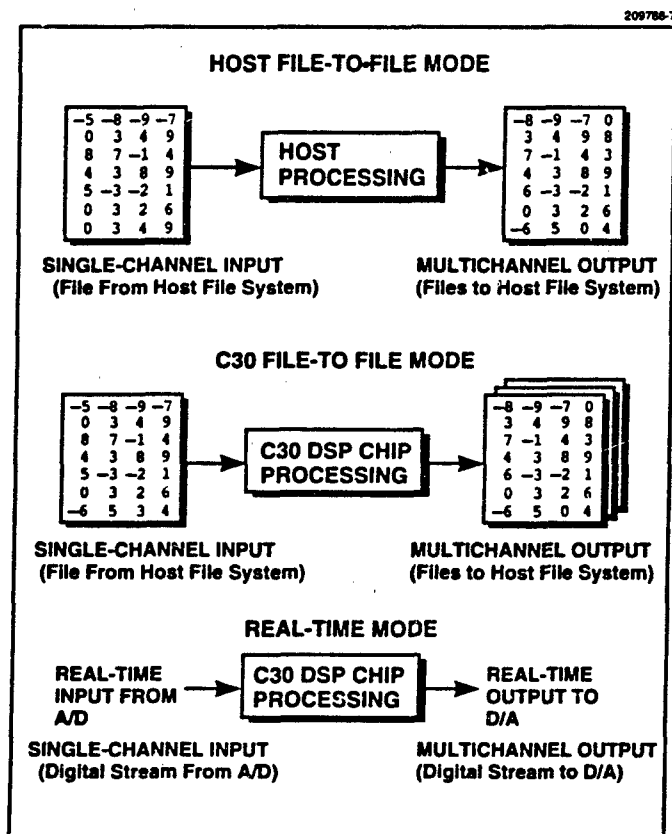


Figure 7. Block diagram of the three operating modes.

3.2.1 Step 1: Host File-to-File Mode

The purpose of a host file-to-file mode is to allow the researcher to explore a range of signal-processing algorithms to be used for electrode stimulation. By processing test files (such as sums of sine waves, noise, chirp signals, and speech), the algorithms can be debugged and evaluated under tightly controlled conditions. All programs are written in C using the standard math and file I/O libraries. Programs are debugged using high-level debugging tools such as Sun's `dbxtool`, an OpenWindows-based, symbolic source code debugger. Finally, the resulting multichannel output files are displayed using interactive waveform and spectrogram display packages such as Entropic Research Laboratory's `waves+`. The final result of step 1 is a debugged C program that takes single-channel speech files as input and produces multichannel electrode stimulation files as output.

3.2.2 Step 2: C30 File-to-File Mode

The next step in the software development process is porting the C code from the Sun host to the Spirit-30 board. An advantage of using the TMS320C30-based Spirit-30 board is the availability of Spectron Microsystems' SPOX operating system. Although TI provides an ANSI C compiler for converting C code to TMS320C30 assembly code, SPOX eases the porting process by allowing the user to retain I/O with the host operating system through the use of commonly used file I/O routines such as `fprintf`, `fwrite`, `fscanf`, and `fread`. In this manner, the file-to-file C program written for the Sun can be recompiled and run on the C30 with very few changes.³ All interaction between the host and the TMS320C30 board required to effect transfer of data from disk to TMS320C30 is performed (from the user's point of view) invisibly. Using SPOX and comparing the output of the host file-to-file system with the C30 file-to-file system, the user can identify quickly any compiler or floating-point inconsistencies between the Sun and TI CPUs. Additionally, C30 file-to-file mode eases the transition to real time by allowing the user to convert from C to C30 assembly language those subroutines containing critical loops that are identified using the on-board C30 timer. Both the correctness and the effective speed-up of these conversions to assembly language can be verified using test files as input.

3.2.3 Step 3: Real-Time Mode

For the real-time mode, the C and assembly code tested in the C30 file-to-file mode are extended to support input from and output to the FLAIR board in addition to the host file system. Prior to entering this mode, the user has an in-hand code that has been verified correct in the C30 file-to-file mode; thus, debugging attention can be focused on porting to the analog interface. C-callable assembly language subroutines are available to the programmer for easy setup of the analog

³However, careful attention to memory management is required. Programmers are advised to use SPOX memory allocation routines, which differ slightly from traditional C.

interface parameters (sampling rate, number of channels, and anti-imaging filter specification), definition of input and output memory buffers, and initiation and termination of analog conversion.

4. ALGORITHMS

4.1 Digital Simulation of the Ineraid Stimulator

The first algorithm implemented on PISCES was a digital simulation of the analog processing performed by the Ineraid stimulator as shown in Figure 8. The signal is first fed to an automatic

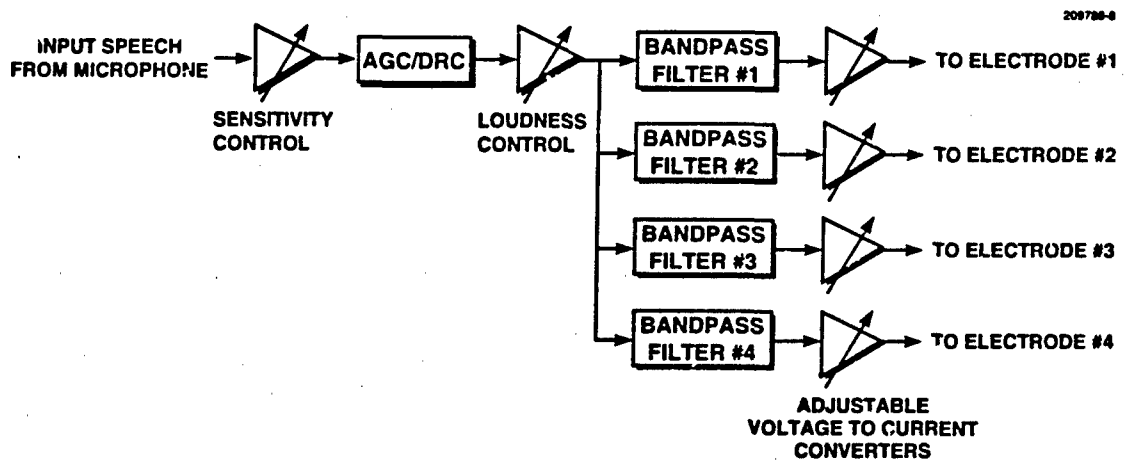


Figure 8. Signal processing performed by the Ineraid hardware stimulator and the digital simulation.

gain control (AGC) which applies dynamic range compression (DRC). The output of the AGC/DRC is sent to a four-channel filter bank. In the analog hardware, each resulting continuous waveform output from the filter bank drives a voltage-to-current converter whose gain is adjusted to reflect the threshold measured for the subject's corresponding electrode. In the digital implementation, the outputs of the filter bank are sent to a D/A converter after which output gain and voltage-to-current conversion are applied by external hardware. In both cases, the output of the voltage-to-current converters are used to stimulate the Ineraid implant electrodes.

The main advantage of the digital simulation over the analog implementation is the degree of flexibility afforded to the clinical staff in fitting a stimulator to an individual subject. In the analog implementation only gain parameters may be adjusted. In the digital simulation, essentially all of the AGC/DRC and filter bank characteristics may be specified at run time. The following sections summarize the digital simulation software, demonstrate its flexibility, and show example input and output signals for a typical configuration.

4.1.1 The cbank Program

cbank is a C language digital simulation of the Ineraid stimulator. The program has been compiled and run successfully on both a Sun SPARCstation IPC and the Sonitech TMS320C30 board. For the C30 board, a few key subroutines were optimized in assembly language to achieve real-time performance.

cbank was written and tested using the three-step procedure outlined in Section 3.2. When running on the Sun, **cbank** reads a single-channel sampled data file from the host file system and produces a multichannel file as output. When running on the C30 board, the user may instruct **cbank** to either read and write files from/to the host file system through SPOX or read and write from/to the A/D and D/A converters.

Considerable flexibility is available to the user through the use of command line arguments and specification files. These arguments and files allow the user to configure a wide range of **cbank** parameters at run time, rather than at compile time, thereby enhancing the user's ability to work with a subject interactively. Parameters that can be varied include sampling rate, number of filters in the filter bank, filter shapes, AGC/DRC characteristics, I/O type (disk versus A/D and D/A), etc. Table 1 and Figures 9 and 10 show the complete list of command line arguments, an example main parameter specification file, and an example DRC file, respectively. The parameter specification file of Figure 9 identifies four files each containing coefficients for an FIR filter. These four files are read during program initialization, and the corresponding coefficients are stored in memory. In addition, the use of AGC, as well as the AGC attack and decay time constants, is specified.

4.1.2 Automatic Gain Control/Dynamic Range Compression

In the digital simulation, the AGC/DRC is implemented as follows:

- For each input sample, determine whether the AGC is in attack mode or release mode.
- Depending on the mode, calculate a new estimate of the signal envelope.
- Given the signal envelope estimate, calculate an appropriate gain and apply this gain to the sample.

Defining $x[n]$ as the output of the A/D converter at time n , the attack mode envelope is defined as

$$y_A[n] = \alpha_A |x[n]| + \beta_A y[n-1] \quad (1)$$

The corresponding release mode envelope is defined as

$$y_R[n] = \alpha_R |x[n]| + \beta_R y[n-1] \quad (2)$$

TABLE 1
The cbank Command Line Arguments

Flag	Value Type	Description	Default
	String	Input speech file	(No default)
	String	Output stimulation file	(No default)
-bs	Integer	Input buffer size (samples)	50
-mxf	Integer	Maximum number of filters	20
-sr	Integer	Sampling rate in Hz	10000
-sf	String	Main parameter specification file	"specfile"
-df	String	DRC specification file	"testdrc"
-v	Integer	Diagnostic Verbosity level	0
-tc	(None)	Enable timing check using C30 timer	FALSE
-tr	Integer	Timer resolution in us	100
-aio	(None)	Use analog interface I/O (AIO) instead of file system I/O	FALSE
-ib	Integer	AIO: number of input buffers	2
-ob	Integer	AIO: number of output buffers	2
-rtm	(None)	AIO: enable real-time modification of parameters	FALSE
-rtmup	Integer	AIO: real-time update time (seconds)	3

If $x[n] > y[n-1]$, then the AGC is defined to be in attack mode, and $y[n]$ is set equal to $y_A[n]$. On the other hand, if $x[n] \leq y[n-1]$, then the AGC is defined to be in release mode, and $y[n]$ is set equal to $y_R[n]$. Once the mode is determined and $y[n]$ is calculated, $y[n]$ is used as an index to a lookup table to determine $g[n]$, the gain for time n . The final output of the AGC/DRC, which is used as the input to the filter bank, is

$$z[n] = x[n] \times g[n] \quad (3)$$

The α and β values are derived from the attack and release times, t_A and t_R , set in the main parameter specification file. Given an A/D sampling period τ in seconds, and t_A and t_R specified by the user in milliseconds,

$$\beta_A = e^{-1000\tau/t_A} \quad ; \quad (4)$$

$$\beta_R = e^{-1000\tau/t_R} \quad ; \quad (5)$$

209788-9

```

VERSION=22
BEGIN_FRONTEND
# specify gain prior to AGC in dB
pregain=0
END_FRONTEND
BEGIN_FILTERS
# format: filter-file-name post-gain-to-be-applied-dB
filters/0000.0800.dat -15.4
filters/0700.1300.dat -9.4
filters/1300.2400.dat -4.5
filters/2300.4400.dat -0.0
END_FILTERS
BEGIN_AGC
# values in milliseconds. first line either "enabled" or "disabled"
enabled
attack=0
release=250
END_AGC

```

Figure 9. A cbank main parameter specification file.

209788-10

```

-100 -100 -50 -50
-50 -50 -40 -10
-40 -10 0 -6

```

Figure 10. A cbank dynamic range compression file.

$$\alpha_A = 1 - \beta_A \quad ; \quad (6)$$

$$\alpha_R = 1 - \beta_R \quad . \quad (7)$$

The DRC function $g[\]$ is provided indirectly by the user to `cbank` in a DRC specification file as a piecewise linear function of desired output envelopes e as a function of estimated input envelopes y . For convenience, the user specifies this function by providing the endpoints (in decibels) of each linear component. The linear gains g are derived by the linear domain division of e by y and are calculated as a function of y during algorithm initialization. Three piecewise linear regions have been specified in the example DRC specification file of Figure 10. This specification, shown graphically in Figure 11, is linear for low envelope levels, expansive for a narrow range of intermediate envelope levels, and compressive for normal envelope values.

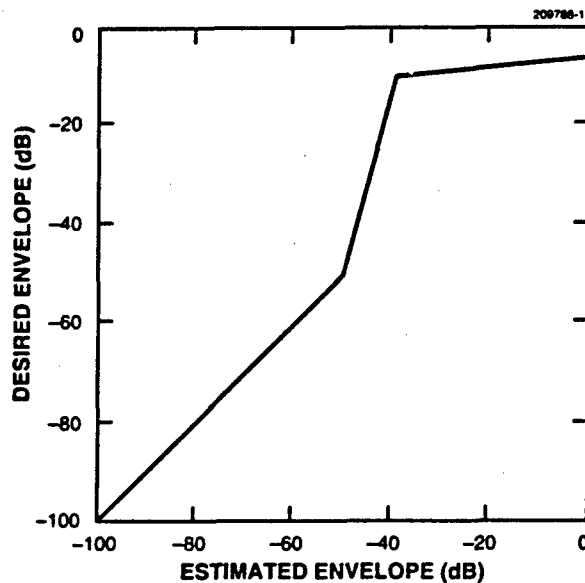


Figure 11. Example of dynamic range compression curve.

4.1.3 Filter Bank Design

Finite impulse response (FIR) filters were employed exclusively in this digital simulation of the Ineraid stimulator for three reasons. First, FIR filters have a distortion-free, linear phase. Second, most signal-processing chips can be programmed easily to perform FIR filtering at the rate

of one tap per cycle with very low overhead. Finally, when performing low-pass filtering followed by downsampling, FIR filter outputs may be computed at the more efficient downsampled rate. The Parks-McClellan procedure was used for designing sharp transition, rectangular frequency response, band-pass filters [9]. The Kaiser window procedure was used to design nonrectangular filters with responses more similar to those obtained with analog discrete components [6].

4.1.4 Waveform Examples from cbank

For the band-pass filter bank shown in Figure 12 and the main and DRC specification files of Figures 9 and 10, Figure 13 shows the output of cbank for a typical speech input. Band-pass filter outputs such as these were used to drive isolated voltage/current converters that provided the stimulation currents for four Ineraid implant electrodes.

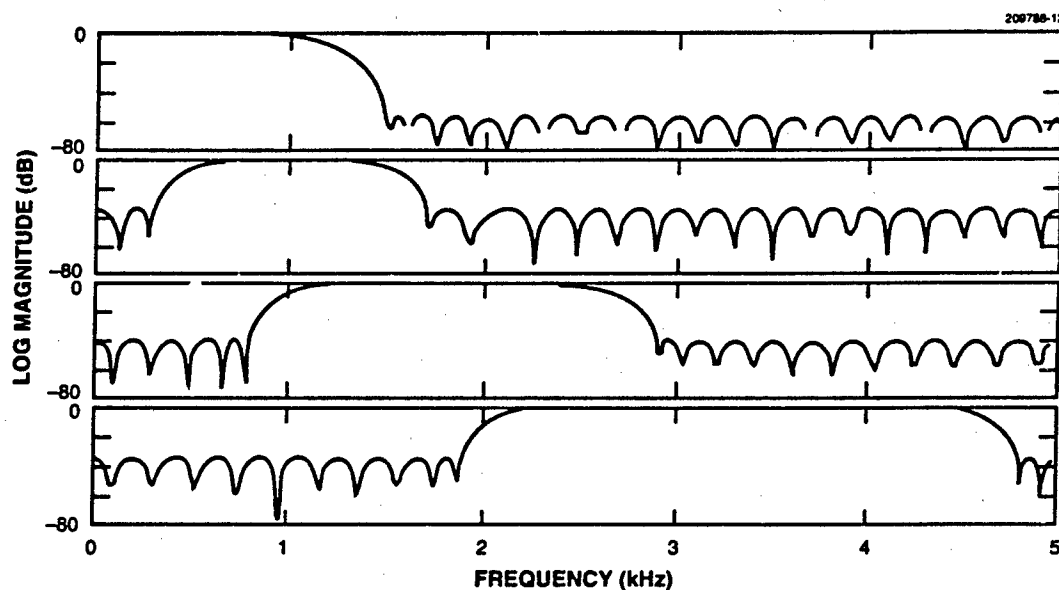


Figure 12. Example of cbank band-pass filters.

4.1.5 Six-Channel Simulations

Present Ineraid implant assemblies provide six signal electrodes and a common ground return. Up to now, the Ineraid analog hardware stimulator has been capable of driving only four signal

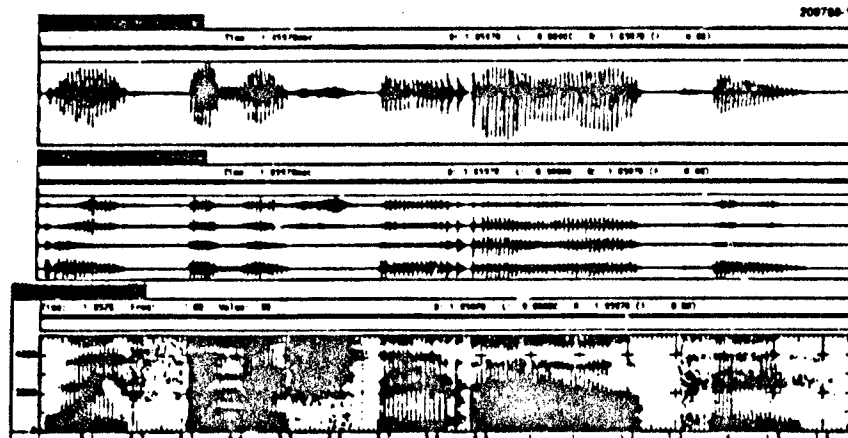


Figure 13. cbank input and output. Top figure shows waveform for input sentence: "We finished the IRP." Middle figure shows four channels of corresponding cbank output. Bottom figure shows a wideband spectrogram of the input.

electrodes. Because the cbank Ineraid simulation is capable of driving six channels, it is possible to run experiments using the full capability of the Ineraid implant.

4.2 Continuous Interleaved Sampling Algorithm

At least two aspects of the Ineraid stimulator design contribute to degraded performance. When the stimulation currents are simultaneously applied from a group of band-pass filters, there is an electric field overlap between the implanted electrodes that results in crosstalk. Second, performance is degraded due to a lack of dynamic range control for each band-pass filter output. Aside from the general effect upon the dynamic range of the entire input spectrum controlled by the input AGC processing, there is no mechanism for fitting the dynamic range of the filter outputs to the useful perceptual range, from threshold to maximum acceptable loudness, of each driven electrode. Output gain adjustments for each channel only reflect the subject's detection threshold.

The Continuous Interleaved Sampling (CIS) stimulator shown in Figure 14 is an attempt to overcome both shortcomings of the Ineraid stimulator [12]. Instead of stimulating electrodes with continuous outputs from a bank of band-pass filters, the CIS stimulation outputs are pulse trains that are amplitude modulated by the output envelopes from a bank of band-pass filters. As shown in Figure 15, the pulse outputs are skewed in time so that electrodes do not receive simultaneous stimulation. Rather than using the raw band-pass filter output envelope to modulate the pulse waveforms, parallel compressors map the raw envelope values into the dynamic range of the subject's individual electrodes. Thus, while the Ineraid and CIS stimulators both use a bank of band-pass filters to generate electrode stimulations, the CIS stimulator transforms the continuous

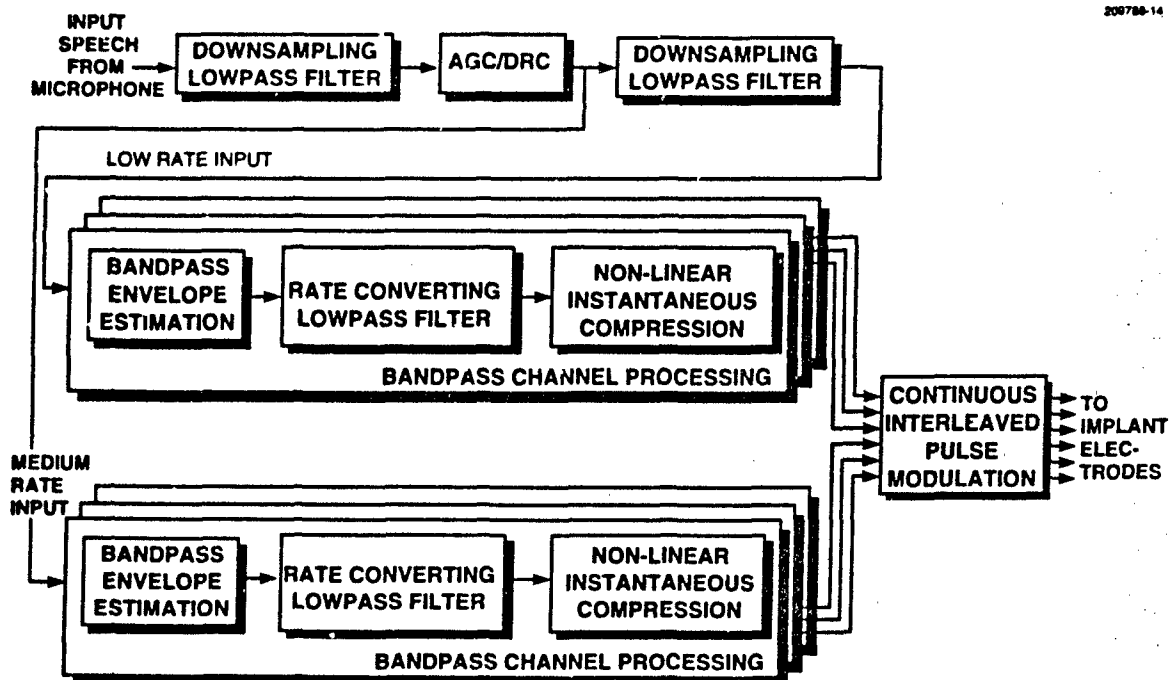


Figure 14. The continuous interleaved sampling stimulator.

filter outputs into envelope modulated, nonoverlapping pulses. At the present time, this CIS design is implemented only on PISCES and does not yet exist as a wearable stimulator for general subject use.

4.2.1 The pbank Program

pbank is a C language program that implements the CIS algorithm. As in the case of the **cbank** program described in Section 4.1.1, the program has been compiled and run successfully on a Sun SPARCstation IPC as well as on the Sonitech TMS320C30 board. A few key subroutines were assembly language optimized to achieve real-time performance.

When running on the SPARCstation, **pbank** reads a single-channel, sampled data file from the host file system and produces a single multichannel file as output. When running on the C30 boards, the user may instruct **pbank** to either read and write files from/to the host file system through SPOX or read and write from/to the A/D and D/A converters.

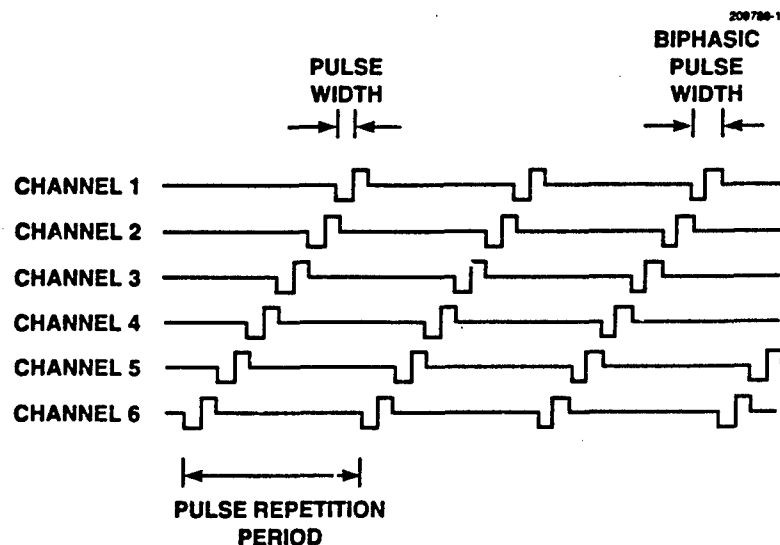


Figure 15. Pulse outputs from pbank.

Command line arguments and specification files allow each of the CIS processing blocks to be altered as required by clinical interactions and experiments. Although the AGC operation is described by the same DRC file format as in the `cbank` program, a more complex main specification file is used. The output compression curves are specified as a series of line segments as in the DRC file or in a table. Table 2 and Figure 16 show the `pbank` command line argument list and an example of a main specification file, respectively. In the following sections each of the processing blocks is described in some detail. Like `cbank`, `pbank` was also designed to provide an arbitrary number of stimulation channels, making it possible to test both four- and six-channel implementations.

4.2. Input Filtering and AGC/DRC

The input processing blocks consist of two low-pass filters and the AGC. The first low-pass filter permits downsampling of the basic system sampling rate. This rate is set fairly high to have a narrow output pulse width from the D/A converter (e.g., a sampling frequency of 32 kHz allows the D/A converter output to produce a minimum pulse width of 31.25 μ s). Because the input spectrum of interest is only 0 to 8 kHz, more efficient signal processing can be employed by downsampling to 16 kHz.

AGC/DRC is exactly the same process described for the `cbank` program and operates upon the output of the first downsampling filter to reduce the overall input dynamic range as specified

TABLE 2
The pbank Command Line Arguments

Flag	Value Type	Description	Default
	String	Input speech file	(No default)
	String	Output stimulation file	(No default)
-bs	Integer	Input buffer size (samples)	96
-mx	Integer	Maximum number of channels	20
-sr	Integer	Sampling rate in Hz	48000
-sf	String	Main parameter specification file	"specfile"
-df	String	DRC specification file	"testdrc"
-v	Integer	Diagnostic verbosity level	0
-tc	(None)	Enable timing check using C30 timer	FALSE
-tr	Integer	Timer resolution in us	100
-hwave		Use half-wave rectification	FALSE
-nohilb		Use only one BPF per channel	FALSE
-nocomp		No output compression	FALSE
-lcomp		Linear interp output compression	FALSE
-tcomp		Table lookup output compression	FALSE
-oclen	Integer	number of pts (li:segs+1 tl:log2(tabentries) in out compression	32
-ocname	String	output compression specification file	"octable"
-aio	(None)	Use analog interface I/O (AIO) instead of file system I/O	FALSE
-ib	Integer	AIO: number of input buffers	2
-ob	Integer	AIO: number of output buffers	2
-rtm	(None)	AIO: enable real-time modification of parameters	FALSE
-rtmup	Integer	AIO: real-time update time (seconds)	3
-hware		AIO: hardware control mode	FALSE
-noup		AIO: no upsampling	FALSE
-fename1	String	Single channel front-ended #1 output file	(No default)
-agcname	String	Single channel AGC'ed output file	(No default)
-fename2	String	Single channel front-ended #2 output file	(No default)
-ename	String	Multichannel envelope output file	(No default)

data-mos1/pbank/less_overlap/spec_hilb_tr

Mon Jul 13 09:37:41 1992

200780-16

1

```
VERSION=1
BEGIN_FRONTEND
# format: filter-file-name post-gain-to-be-applied-dB downsampling-n:1
../more_atten/unity@16000.lpf.dat 0.0 1
../more_atten/2k@16000.lpf.dat 0.0 4
# all channels go thru fir #1. which go thru fir #2?
# legend: 1:thru,0:skip
# 1-3 thru fir#2, 4-6 skip fir#2
1 1 1 0 0
END_FRONTEND
BEGIN_AGC
# values in milliseconds. first line either "enabled" or "disabled"
disabled
attack=0
release=250
END_AGC
BEGIN_ENVELOPE
# format: filter-file-name post-gain-to-be-applied-dB downsampling-n:1
# expecting pairs of filters, i.e. one hilbert xform pair per channel
c1@4000.cos.dat 0.0 1
c1@4000.sin.dat 0.0 1
c2@4000.cos.dat 0.0 1
c2@4000.sin.dat 0.0 1
c3@4000.cos.dat 0.0 1
c3@4000.sin.dat 0.0 1
c4@16000.cos.dat 0.0 4
c4@16000.sin.dat 0.0 4
c5@16000.cos.dat 0.0 4
c5@16000.sin.dat 0.0 4
c6@16000.cos.dat 0.0 2
c6@16000.sin.dat 0.0 2
END_ENVELOPE
BEGIN_DELAY
# the delay that we want to add between the hilbert transform and the
# low-pass filter in each channel. This number is specified in samples;
# the actual time in seconds depends on the sampling rate at the output
# of the hilbert transform. Typically, at least one of the values is
# zero; otherwise, we'd be adding artificial delay.
0 5 7 17 35 38
END_DELAY
BEGIN_SMOOTH
# format: filter-file-name post-gain-to-be-applied-dB downsampling-n:1
# if downsampling is negative, it means upsampling
# gains on output started at 12,12,6,0,0,0; now adjusted for peak output
../more_atten/400@4000.lpf.dat 12.0 3
../more_atten/400@4000.lpf.dat 12.0 3
../more_atten/400@4000.lpf.dat 12.0 3
../more_atten/400@4000.lpf.dat 12.0 3
../more_atten/400@4000.lpf.dat 12.0 3
../more_atten/400@4000.lpf.dat 12.0 6
END_SMOOTH
BEGIN_PM
# pulse modulation pattern specified as floats.
# one column per channel. when we get to the bottom, jump to the top.
extent=12
0 0 0 0 0 0 -1
0 0 0 0 0 0 1
0 0 0 0 -1 0 0
0 0 0 0 1 0 0
0 0 0 -1 0 0 0
0 0 0 1 0 0 0
0 0 -1 0 0 0 0
0 0 1 0 0 0 0
0 -1 0 0 0 0 0
0 1 0 0 0 0 0
-1 0 0 0 0 0 0
1 0 0 0 0 0 0
END_PM
```

Figure 16. A pbank main parameter specification file.

by the attack and release time constants and the DRC file. The second low-pass filter is used to reduce the computation for the subset of channels whose highest frequency is below 2 kHz by allowing a second downsampling for that subset to 4 kHz. Figure 14 shows three upper frequency channels driven from the first downsampling filter, and three lower frequency channels driven from the second downsampling filter.

4.2.3 Band-Pass Envelope Estimation

Band-pass envelope estimation is required to provide a modulating signal to the pulse train outputs. Two envelope estimation methods were implemented. The *rectification* (or detection) method comprises a band-pass filter followed by a strong nonlinearity such as a full-wave or half-wave rectifier. The nonlinearity output is smoothed by a low-pass filter to eliminate spurious harmonics. The quadrature estimation method calls for creating a second band-pass output which is shifted 90° in phase from the original. These two signals are squared, summed, and square rooted, producing an estimate of the envelope. The choice of estimation procedure is specified on the command line. The rectification estimation method is used commonly in both analog and digital systems. Unfortunately, full- or half-wave rectification generates a range of spurious harmonics that alias in the sampled data domain. Appendix A shows that the quadrature estimation method is somewhat more robust to harmonic distortion and, consequently, aliasing.

The filter bank divides the input spectrum into six channels spaced logarithmically in center frequency and bandwidth over the range from 300 to 7000 Hz. Rectification envelope estimation requires only one band-pass filter per channel. Either of the techniques described for the cbank filter bank, namely, Parks-McClellan or Kaiser window design, provides the needed flexibility. Quadrature envelope estimation requires one quadrature pair of band-pass filters per channel. Two approaches for designing such filter pairs have been studied. In the first approach, a prototype low-pass filter is designed using the Parks-McClellan algorithm. Two band-pass impulse responses are obtained by multiplying the low-pass impulse response by sampled sine and cosine functions whose frequencies are at the desired band-pass filter center frequency. Multiplication by sine and cosine guarantees the fixed 90° phase difference [10]. A second design technique uses the eigenfilter method developed by Nguyen which approximates arbitrary magnitude and phase responses in a minimum mean square error sense [8]. This technique was used to generate quadrature filter pairs with analoglike 12 dB per octave responses. In this design procedure, there are small differences in the magnitude response for a band-pass pair in the frequency range of interest. Presently, only the Parks-McClellan frequency shifted filters have been used with subjects.

4.2.4 Low-Pass Smoothing Filters

For the case of a full- or half-wave rectified band-pass filter output, the low-pass filtering process eliminates the out-of-band harmonics that have been generated by the nonlinearity but allows the basic envelope waveform to be passed. This function is not needed for the quadrature derived envelope as discussed in Appendix A. As this smoothed envelope will be the only information provided to the subject, there is a trade-off between wide bandwidths that allow for maximal

envelope variations and narrow bandwidths that reduce pitch harmonic ripple. Additionally, low-pass filtering allows for down- or upsampling so that envelope samples are generated at the pulse sampling rate. At present, the low-pass filter cutoff response resembles a low-order analog filter with a cutoff frequency at about 400 Hz.

4.2.5 Output Compression Mapping

The envelope waveform dynamic range is mapped into the measured dynamic range for the corresponding electrode. The mapping curve is one of the variables to be determined from clinical interaction. Typical transformations that map from x (the envelope estimate available from the low-pass smoothing filter) to y (the modulation level) are given by

$$y = A + B(x - X)^p \quad , \quad (8)$$

where A , B , X , and p are dependent on subject threshold and dynamic range measurements and the desired compression characteristic. This mapping is specified independently for each channel.

4.2.6 Pulse Modulation Waveform Output

The final pulse output waveforms for each channel are computed at the full input sampling rate. For example, a 32-kHz rate will allow a channel D/A to output a pulse of width 31.25 μ s. The specification file of Figure 16 defines the pulse sequencing as a matrix of channels versus activity at each sampling time for one period of the output cycle. The matrix shown in Figure 16 at a sampling rate of 32 kHz would produce the six-channel output of Figure 15, with a pulse repetition period of 375 μ s, pulse width of 31.25 μ s, and biphasic pulse width of 62.5 μ s. The biphasic pulse shape allows for a zero mean output signal⁴ while retaining a narrow pulse shape; as discussed earlier, the skewed, nonoverlapping pulse waveforms eliminate field overlap between electrodes in the cochlea. These modulated pulse trains are output to the voltage-to-current converters, which in turn stimulate the corresponding electrodes.

4.2.7 Waveform Examples from pbank

Figure 17 shows the output of pbank for a typical speech input.

⁴Each electrode is stimulated by a current that is proportional to the pbank waveform output. A zero mean signal results in delivery of zero net charge by the electrode to the cochlea, thereby causing minimal trauma in the surrounding cochlear tissue.

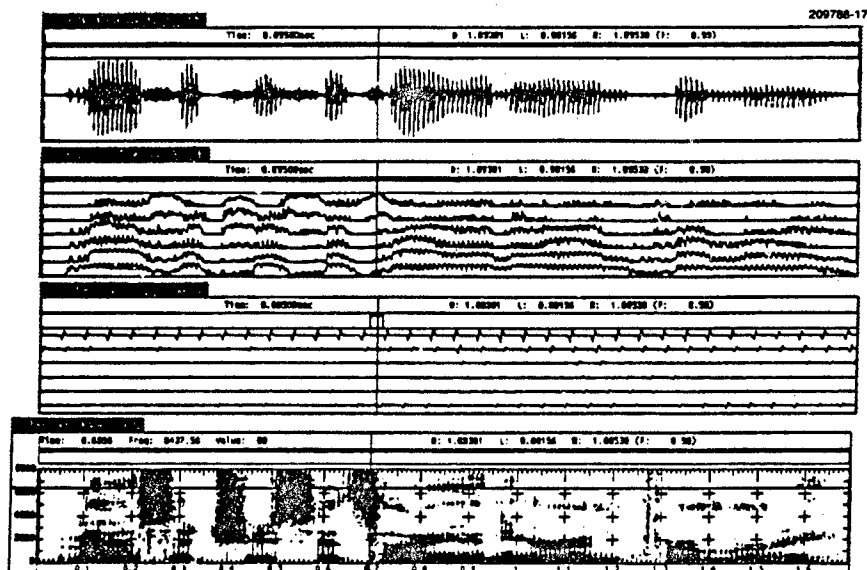


Figure 17. pbank input and output. Top figure shows a waveform for input sentence: "Massachusetts Eye and Ear Infirmary." Next figure shows estimated envelopes for all six channels. Third figure shows a close-up of the pulse output in the middle of an unvoiced fricative. Bottom figure shows wideband spectrogram of input.

5. CLINICAL EXPERIENCES AND RESULTS

The PISCES system was installed at the Massachusetts Eye and Ear Infirmary's (MEEI) Cochlear Implant Research Laboratory (CIRL) in early July 1991. The output of each D/A channel of the audio interface was connected to an isolated current stimulator whose gain is adjustable independently and whose output level is monitored and limited to a preset maximum value. This set of isolated current drivers and monitoring circuits is contained in a single equipment rack outside of a small sound-insulated testing room. The current outputs are available on a cable/plug assembly inside the testing room so that subjects can unplug their own stimulator and substitute the isolated output currents driven by PISCES. In addition, a "panic button" accessible to the subject allows for a rapid disconnect between the connector and the current drivers in an emergency. A photograph of PISCES and the current isolator/limiter equipment rack is shown in Figure 18.

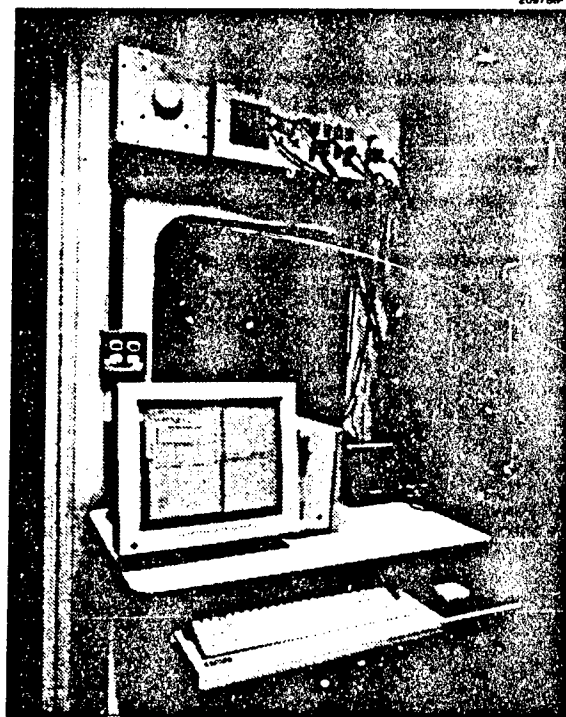


Figure 18. Photograph of PISCES and current isolator/limiter equipment at MEEI CIRL.

The first clinical interaction presented subject S04 with the outputs of PISCES running **cbank**. Initially, subject S04 described the **cbank** sound processing as very different from his own Ineraid hardware stimulator. However, after adjusting the four band-pass filters to approximate more closely the analog hardware filters, the **cbank** stimulations were judged as being very similar to the Ineraid hardware stimulations. This interaction assured us that PISCES and the current isolator/limiter equipment were capable of replacing safely and adequately the Ineraid hardware stimulator. All remaining clinical tests focused on CIS algorithms as implemented in **pbank**.

5.1 Psychophysical Measurements

To specify the output compressions of the **pbank** system, the dynamic range of each of the implant electrodes must be determined. Dynamic range is defined as the difference between the signal intensity at threshold (i.e., a just perceivable signal) and the signal intensity that is just uncomfortably loud (UCL). This represents the range over which stimulating signals can be usefully applied and perceived. The range is a measure of the physiological state at the site of electrode action and is a function of the induced current's proximity to still functioning auditory nerves. Not only do dynamic ranges vary among electrodes but they are also functions of the pulse widths and frequencies used for the stimulus signal. As a consequence, it is important that the dynamic range be measured for the pulse parameters of the stimulator in use.

The dynamic range measurements are used to set the output gains and/or compression characteristics of the stimulating currents. In the case of the Ineraid hardware stimulator, the subject's threshold measurements are used to set the gains for each filter output. For **pbank**, the dynamic range information defines the output range of the compression curve for each channel. Figure 19 shows a typical output compression curve for one electrode. The peak value observed from the envelope estimator is always mapped to UCL. The input dynamic range, defined as the distance between the envelope outputs mapped to threshold and mapped to UCL, is adjusted by moving the low cutoff value.

5.2 Speech Materials

When a subject is initially connected to PISCES running **pbank**, an informal interactive conversation between the subject and the researcher is used to gauge gross performance. This provides useful feedback about signal levels as well as crude comparisons between the present and previous parameter settings. In addition, the subject has the opportunity to acclimate to each new stimulator variation, thereby providing at least a small amount of learning before more quantitative testing takes place. In addition, the subject may identify crude bugs in the new stimulation system under test. During this phase of testing, fine tuning of the isolator gains may be performed.

The main quantitative tests of speech reception were based on measures of consonant identification. These tests make use of 24 consonants in a vowel-consonant-vowel (VCV) setting (e.g., "asha," "aba") spoken by a male talker and available on a laser videodisc from the University of Iowa [11]. An IBM-PC/AT control program written specially for the videodisc database (and

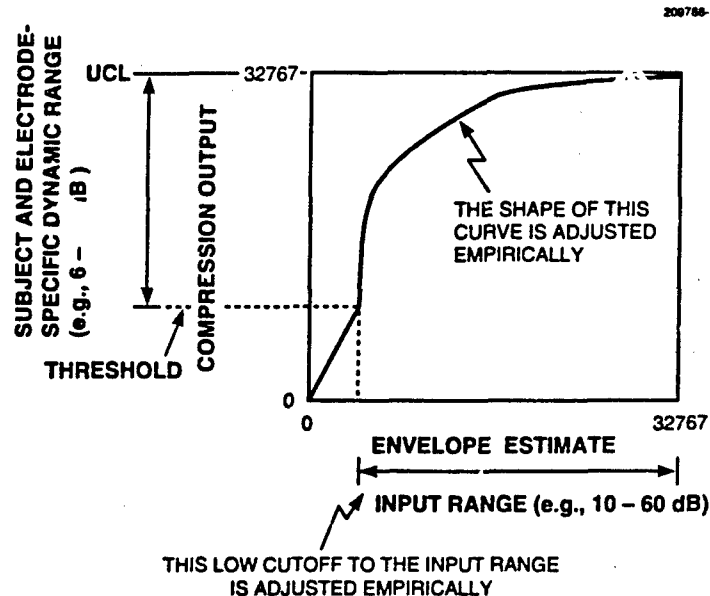


Figure 19. A typical output compression curve.

provided by researchers at the Research Triangle Institute (RTI) plays random sequences of VCV utterances and tabulates the subjects' responses [13]. The researcher may choose to use subsets of 8 or 16 consonants or may choose to use the full 24 consonant set. In all cases, 5 groups of the 8, 16, or 24 randomized consonant sets are presented. The use of this consonant test and test controller program allowed direct comparison with RTI results.

5.3 Subject Experiments

Six subjects were connected to PISCES running *pbank* as described below.

5.3.1 Experience with Subject S04

Subject S04 appears to have good nerve survival, as shown by the measured dynamic ranges for the six monopolar electrodes of the Ineraid implant in Table 3. These measurements were made for 200-Hz biphasic pulse trains with a 500- μ s biphasic pulse width. Table 4 shows the results of the 24 consonant VCV tests on S04. As a baseline, his score using his Ineraid hardware stimulator averaged 79% correct over six separate tests. The interaction with S04 using PISCES and *pbank*, which spanned over 40 h of testing, was aimed at substantially raising his speech reception score by adjusting *pbank* parameters. The clinical interactions began by adjusting *pbank*'s channel specific

TABLE 3**Dynamic Range Measurements for Subjects S04 and S05**

Subject S04						
Electrode Number	1	2	3	4	5	6
Threshold (μ A)	32	31	33	38	40	67
UCL (μ A)	210	235	285	280	305	300
Dynamic Range (dB)	16.3	17.7	18.8	17.3	17.7	13.0
Subject S05						
Threshold (μ A)	73	66	71	134	116	90
UCL (μ A)	162	177	202	250	187	163
Dynamic Range (dB)	6.9	8.6	9.1	5.4	4.2	5.2
Dynamic Range Differences, S04-S05						
Difference	9.4	9.1	9.7	11.9	13.5	7.8

TABLE 4**Subject S04 Performance on 24 VCV Test**

Condition	Variable	Scores
Ineraid Hardware	Baseline	79% (73-85)
6-Channel CIS	12 dB/oct. BPFs	87% (83-92)
FWR envelopes	Rectangular BPFs	96% (95-96)
6-Channel CIS	FWR envelopes	96% (95-96)
Rectangular BPFs	Quadrature envelopes	99% (98-99)
6-Channel CIS	Interleaved pulses	99% (98-99)
Rectangular BPFs	Coincident pulses	82% (one test)
Quadrature envelopes		
CIS	6 Channels	99% (98-99)
Rectangular BPFs	4 Channels	89% (one test)
Quadrature envelopes		

output compression curves to match the measured values of S04's dynamic range. As these dynamic ranges are dependent upon the pulse widths and pulse rates used for the stimulating signals, the compression curves were recomputed whenever these parameters were changed. Generally, mapping a 60-dB envelope dynamic range to a 15- to 18-dB electrode dynamic range was optimal.

Next, two different filter bank designs were evaluated. The first filter bank shown in Figure 20 is an approximation to a bank of analog, second order, Butterworth filters. The second filter bank

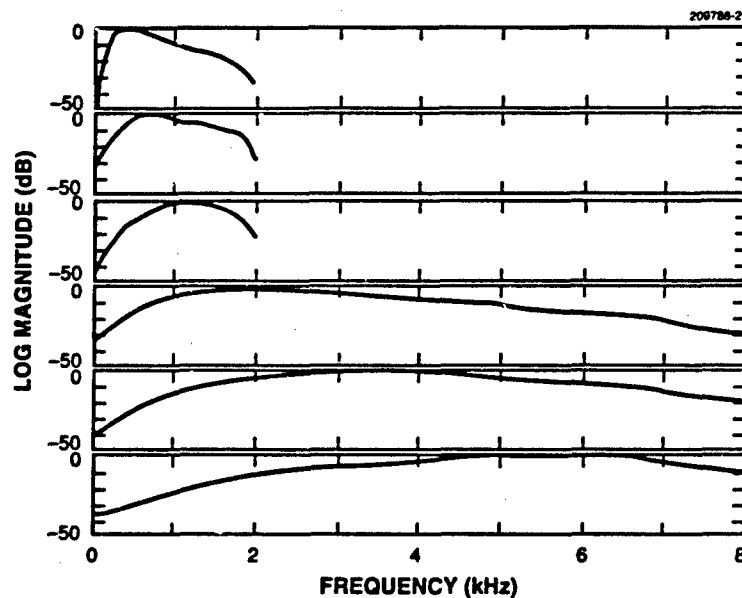


Figure 20. Filter bank approximating Butterworth filters.

shown in Figure 21 is a more rectangular, frequency selective design. For S04, the rectangular filters were superior, and, using these filters, quadrature processing outperformed the full-wave rectification.

Informal experiments varying the pulse repetition rates in the range from 1 to 2 kHz with corresponding pulse widths of 31 and 62 μ s for each subpulse did not show much difference. However, nonoverlapping presentation of each channel's pulse stimulation resulted in superior performance versus coincident presentation, suggesting that reducing interelectrode field interaction is beneficial. Only a single test with coincident pulses was run as it was clear from the subject's comments that these pulses caused a significant degradation and change in the effective loudness of the signals.

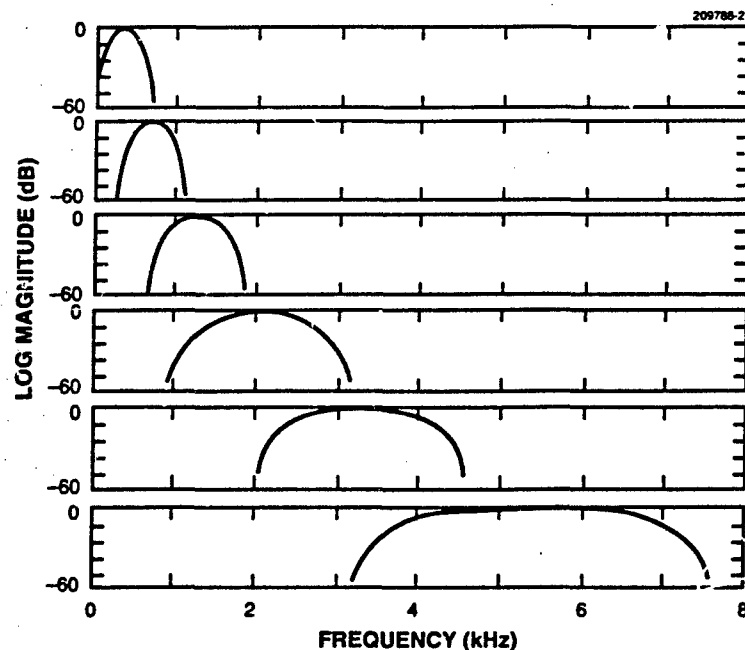


Figure 21. Filter bank with more rectangular filters.

Testing of a four-channel CIS stimulator produced scores 10% less than the six-channel CIS system, indicating that the extra two channels provide added benefit. A four-channel CIS outperformed the Ineraid stimulator by approximately 10%, suggesting that the interleaved stimulation and the output compression also contribute to the superior CIS performance.

The interaction with S04 in the context of PISCES running pbank demonstrated the basic hypothesis of the IRP, namely, that it would be possible to improve the speech reception performance of cochlear implant subjects by modifying stimulator parameters interactively. Because the surgical placement of electrodes and number of surviving neurons varies from subject to subject, it is important to be able to adjust an electrode stimulator for each subject's condition.

It is worth noting that S04 often chose to remain in the clinic for hours listening to music through the PISCES system. Thus, CIS shows promise of improving acoustic reception for a wide range of inputs.

5.3.2 Experience with Subject S05

Table 3 shows that subject S05 has a narrower set of dynamic range measurements than S04, presumably due to poorer nerve survival and electrode placement. The differences between the ranges of S04 and S05 vary from 9 to 13 dB. Since the dynamic ranges are quite small compared to a normal hearing range of over 100 dB, these differences would be expected to yield qualitatively different behavior. As a baseline, S05 scored 34% on the 24 consonant test using his Ineraid stimulator (cf. S04's 79%).

S05's results using CIS are shown in Table 5. The CIS parameter set that produced the best

TABLE 5
Subject S05 Performance on 16 VCV Test

Gain (dB)	Input Dynamic Range (dB)					
	60	40	30	20	10	Linear
-2				49.0%		
0	46.0%	57.0%†				
4				61.2%		
10		55.0%	56.4%	54.8%		
16				63.9%‡	57.8%	55.0%
24						53.0%
0†		49.5%				
16†				56.4%		61.4%
Compare to Ineraid hardware baseline performance of 48.9.†						
†Boost of 3, 6, 9, and 12 dB on channels 3-6, respectively.						
‡Condition for which two test runs were performed.						

scores for S04 (i.e., a system using a 60-dB input range, quadrature envelope extraction, sharp filters, and six channels) resulted in a disappointing result of 23% on the 24 VCV test. This result demonstrates that a parameter set producing good performance in one subject may not be good for another.

As there was no reason to assume any of the parameters of the CIS system that worked best for S04 should not be a good starting point for S05 (the only exception being the subject dependent compression mappings for each channel), various sets of input compression ranges and gains were explored. Careful input gain adjustment is required to maximize the speech activity within the

narrow compression range. Testing was reduced to only 16 consonants (a subset of the 24) to keep S05 from tiring as a result of the difficult (for him) 24 consonant test. Table 5 shows the result of the dynamic range exploration and again points out the importance of tuning the stimulator for each subject's individual condition. It appears that S05 cannot process the wide dynamic range inputs that S04 finds most usable and pleasant. In fact, when the input range available to each of S05's electrodes was restricted, his performance increased significantly. Notice that the best simulation system used for S05 provides an average score for two test runs of 63.9% correct versus an average of two test runs for the Ineraid stimulator of 48.9% correct.

The only parameters adjusted for S05 were the input dynamic range of the compression curve in each channel output, the overall gain, and the per channel gain. Even this simple set of parameter adjustments generates a large space of possibilities that requires many hours of subject interaction. If this space were explored in greater detail, the results might still be a strong function of other stimulator variables (e.g., the band-pass filter responses). A strategy is needed that allows convergence to an overall optimal set of parameters for each subject. In addition, the scores for other sets of speech tests must be examined as well.

5.3.3 Brief Experiences with Four Subjects

Four additional subjects were investigated briefly, each in a single session of approximately 3 h. The results of these sessions as well as the best scores for S04 and S05 are shown in Table 6. Each of

TABLE 6
Best Performance for Six Subjects

Subject	Number of Test Consonants	Ineraid Hardware Score	Best PISCES Score	Comments
S04	24	79% (73-85)	99% (98-99)	50 h of testing
S05	16	49% (44-54)	64% (61-66)	10 h of testing
S02	24	29% (27-31)	33% (30-35)	3 h, chan 6 unused
S01	24	30% (one test)	11% (one test)	3 h, 5&6 unused
S16	8	65% (54-71)	69% (58-75)	3 h
S23	16	38% (25-50)	36% (31-44)	3 h

these subjects was tested with the stimulator found to be optimum for subject S04. Notwithstanding whether S04's parameter set may be a good starting point for the other four subjects, it also became

clear that a single, short clinical interaction is not sufficient to tune a stimulator for a subject's particular needs.

Experience with Subject S02: Testing S02 on the 24 consonant test through his Ineraid stimulator produced an average score of 29.1% for two separate test runs. After computing output compression curves based on dynamic range measurements and setting current isolator gains, the subject observed that channel 6 caused uncomfortable sensations of feeling rather than sound. As a result, channel 6 was not included in the CIS system for this subject.

Two separate test runs, each presenting the 24 consonant set through the CIS stimulator, resulted in an average score of 32.5%. This difference in performance relative to the Ineraid stimulator was small. In fact, S02 preferred his own Ineraid stimulator. In future sessions, the plan is to adjust input compression ranges as done for S05 and to adjust the band-pass filter ranges to compensate for the lack of stimulation from channel 6. Future sessions with S02 might also focus on optimizing a four-channel CIS simulation before trying to integrate the fifth channel.

Experience with Subject S01: S01 scored 30% on the 24 consonant test using his Ineraid stimulator. When electrodes 5 and 6 (unused by his Ineraid hardware stimulator) were stimulated to measure dynamic range, he had great difficulty making judgments of loudness and perceived sensation. S01 has been deaf for many years, and he found the sensations on electrodes 5 and 6 so unusual that he could not judge whether the current levels were reasonable. As a result, the gains of these two channels were set at arbitrary levels for CIS stimulation. He scored 11% on the 24 VCV test.

S01 presents a distinct challenge in trying to extend stimulation to electrodes 5 and 6. As in the case of S02, it might be worthwhile to start the next session by optimizing a four-channel CIS simulation before moving to five or six stimulation channels.

Experience with Subject S16: For S16, it was possible to set up and use all six electrodes. As data regarding threshold and UCL for each electrode were available only at a wider pulse width of 100 μ s per phase, the subpulse width was increased to 125 μ s, and the pulse repetition rate was lowered to 500 Hz. Testing with S16's Ineraid resulted in a score of 65.3% averaged over three test runs using an eight consonant subset. Because S16 complained that the CIS was driving his electrodes too strongly, the peak currents were reduced to 0.75 of the original values. At these settings, three test runs for the eight consonant test gave a score of 71%. Attenuating the current gains by another 0.75 and running two more tests gave an average of 66.5%. The overall average for the five tests was 69.2%.

The slight increase in consonant score when using CIS may be significant. Again, there are many parameters to be adjusted for S16, including choosing the current isolator gains that best use his dynamic range. The most encouraging result of this session was S16's comment that he is "understanding more than he ever has" of ordinary conversation with the CIS simulation in its present unoptimized version.

Experience with Subject S23: S23 scored 37.6% when using the Ineraid stimulator for the 16 consonant test. Tests of the initial CIS stimulator produced a score of 36.2% using the 16 consonant subset. S23's comment that "the clarity is good" for CIS compared to Ineraid gives us optimism that it will be possible to converge to a set of parameters that yields substantial increases in speech reception compared to the Ineraid.

5.4 Comments on the Clinical Experience

Because clinical interaction using systems running real-time, digital speech algorithms connected to cochlear implants was a new activity for both the Lincoln and MEEI staff, there was a steep learning curve from July until October. By October, it became easier to run automated speech tests, to shift parameters easily, and to set reasonable initial CIS parameters.

None of the subjects described have used the CIS systems for lengths of time comparable to the time they have used their Ineraid stimulators. In fact, except for S04 and S05 who underwent some previous testing at RTI, none of the other subjects had any previous CIS experience before the single testing sessions reported. Speech reception with CIS processing would be expected to improve with additional use and with continued interactive searching for better parameter sets. Several subjects commented that use of a new CIS stimulator in place of the Ineraid hardware would improve their day-to-day speech reception performance.

6. CONCLUSIONS

This report has summarized an Innovative Research Program project aimed at improving cochlear implant stimulators. Drawing upon Lincoln Laboratory expertise in speech coding, digital signal-processing theory and design, and facilities, an interactively adjustable implant stimulator, PISCES, was designed, built, and tested. With this system installed at MEEI/CIRL, speech reception of several implanted subjects was improved. Flexible interaction in a clinical environment also enables new designs to be explored and tested.

A need still exists for a wearable/portable stimulator capable of running the algorithms evaluated on PISCES. Such a device should be easily reprogrammed both with respect to algorithm and parameter set and would provide critical information about long-term learning effects for specific stimulator algorithms. In addition, mechanisms for converging toward optimal parameter sets for a given subject and a given stimulator algorithm must be developed, as even the simplest of algorithms has a very wide space of adjustable parameters. The National Institutes of Health will be funding a three-year program at MIT involving all the authors for extending PISCES in these areas.

Although this was not the first effort to provide an ability to modify cochlear implant stimulators interactively in a clinical environment (both the RTI [13] and Melbourne [1] groups preceded this effort), this effort is unique in employing a laboratory stimulator capable of running floating-point algorithms, supporting a high-level programming language, and providing 16-bit analog I/O. These differences may become more important over the next few years as CIRL continues to utilize PISCES.

APPENDIX A

The Advantages of Quadrature Envelope Estimation

The envelope of an analog signal is often estimated by applying a strong nonlinearity such as a full-wave or half-wave rectifier. The output of the rectifier drives a low-pass smoothing filter, which eliminates spurious harmonics generated by the process. In the sampled data domain, due to aliasing, the spurious harmonics generated by the full-wave or half-wave rectifier cannot (generally) be removed by low-pass filtering. As a consequence of this concern, an envelope estimator has been implemented that uses a pair of band-pass filters identical in frequency response but 90° out of phase ("quadrature" filters). The outputs of these quadrature filters are squared, summed, and square rooted, resulting in a different envelope estimate. The following discussion describes some of the differences between rectification- and quadrature-based envelope estimation.

Consider a single sinusoidal input to a band-pass filter. The output of the band-pass filter is

$$y(t) = A \cos(\omega t) \quad . \quad (A.1)$$

Full-wave rectification yields a signal whose Fourier series expansion is

$$|y(t)| = 2A/\pi + (4A/3\pi) \cos 2\omega t - (4A/15\pi) \cos 4\omega t + (4A/35\pi) \cos 6\omega t + \dots (A.2)$$

This operation has generated a series of even harmonic terms. In the analog domain, these terms would be eliminated with a low-pass smoothing filter operation, and the result would be a term at DC representing only the input amplitude (i.e., the envelope). In the sampled data domain, these harmonics of the input frequency at four and six times the input frequency may produce aliasing of the input frequency in the bandwidth of the low-pass filter. For example, a sinusoid at 2700 Hz will produce a sixth harmonic at a frequency of 16,200 Hz which will be aliased to 200 Hz for a system sampling rate of 16 kHz. From the series expansion, this aliased harmonic would be about 2/35 of the DC term or about 6%. Higher input frequencies could produce aliasing of the fourth harmonic term at even higher levels.

For the envelope produced by the quadrature operations, the two band-pass filter outputs are Hilbert transforms of each other (i.e., their outputs are shifted by 90°), yielding

$$y_1(t) = A \cos(\omega t) \quad \text{and} \quad (A.3)$$

$$y_2(t) = A \sin(\omega t) \quad . \quad (A.4)$$

Next, the sum of squares signal is calculated as

$$y_1^2 + y_2^2 = A^2 \cos^2 \omega t + A^2 \sin^2 \omega t = A^2 \quad . \quad (A.5)$$

The squared term has no spurious frequencies, and the final square root operation produces the constant envelope value A . Notice that a low-pass smoothing operation is not even required for this simple case as no harmonics of the signal are generated.

For a more elaborate case of a two-sinusoid signal in the passband of a band-pass filter, the output (ignoring phase offsets) is of the form

$$y(t) = A \cos \omega t + B \cos(\omega + \Delta)t \quad (\text{A.6})$$

The output of a full-wave rectifier operation is difficult to quantify for even this simple two-sinusoid case, but one can speculate that harmonics of the input frequencies as well as sum and difference products will be produced causing aliasing for the higher frequency inputs and the higher order distortion terms.

Considering the quadrature processing of the two-sinusoid signal, the outputs would be

$$y_1(t) = A \cos \omega t + B \cos(\omega + \Delta)t \quad \text{and} \quad (\text{A.7})$$

$$y_2(t) = A \sin \omega t + B \sin(\omega + \Delta)t \quad (\text{A.8})$$

The sum of the squares of the two band-pass filter outputs becomes

$$y_1^2 + y_2^2 = A^2 + B^2 + 2AB \cos \Delta t \quad (\text{A.9})$$

Notice that this signal contains only constant amplitude terms and a term which represents the difference between the two sinusoids (in voiced speech, this would be the pitch frequency). These terms represent the squared envelope term with no spurious frequencies in need of suppression by a low-pass filter or susceptible to aliasing. Unfortunately, the square root operation required to generate the envelope amplitude does generate spurious harmonics. However, these harmonics are located only at multiples of the difference frequencies (the pitch harmonics), not the original input frequencies. For a general sum of sinusoids as an input to the pair of band-pass filters and the Hilbert envelope process, the squared output will only contain DC terms reflecting the energy of each sinusoid and sinusoidal terms at each of the possible difference frequencies. The process of taking the square root will generate harmonics of these difference frequencies, but not harmonics of the original input signals so that there is a lower probability of aliasing spurious energy back into the low-pass filter range. Note that as the envelope estimate is applied to a nonlinear compression curve before it modulates the output pulse train, even a perfect envelope signal containing pitch harmonics would generate harmonics of the pitch signal.

For the Hilbert envelope case where there are only difference frequencies and some amount of difference frequency harmonics because of the root operation, the envelope can be downsampled for computational savings. If the rooting distortion is ignored, then the highest difference frequency in

any band-pass filter Hilbert envelope output will be the difference between the lowest and highest frequency sinusoids that can be passed by that filter—a difference frequency equal to the passband width. In that case, the envelope waveform can be sampled at twice the bandwidth of the filter. If the output of a band-pass filter was a conventional amplitude modulated signal with the full sidebands fitting into the filter bandwidth, then the modulating signal (the envelope) would be half the bandwidth of the filter. In this case, the envelope waveform could be sampled at a rate equal to the bandwidth. Since one cannot model the band-pass outputs as simple amplitude modulated signals, one must be guided by the two times bandwidth rule. To be somewhat conservative, higher rates have been employed when possible.

For the case of outputs from rectification operations, there is no expectation that such signals are band limited. Conservative design principles would dictate upsampling the band-pass filter outputs before applying the rectification operations, thereby lowering the chances of aliasing harmonics of the band-pass outputs back into the band of interest.

REFERENCES

1. G.M. Clark et al., "Speech processors for auditory prostheses," Final Report on NIH Contract NO1-DC-9-2400 (December 1991).
2. P.B. Denes, *The Speech Chain: The Physics and Biology of Spoken Language*, Garden City, NY: Anchor Press (1973).
3. M.F. Dorman et al., "Acoustic cues for consonant identification by patients who use the Ineraid cochlear implant," *J. Acoust. Soc. Amer.* 88, 2074-2079 (1990).
4. D.K. Eddington, "Speech discrimination in deaf subjects with cochlear implants," *J. Acoust. Soc. Amer.* 68, 885-891 (1980).
5. T. Ifukube, "Signal processing for cochlear implants," in S. Furui and M. M. Sondhi (ed.), *Advances in Speech Signal Processing*, Marcel Dekker (1992), pp. 269-305.
6. J.F. Kaiser, "Nonrecursive digital filter design using the I_0 -sinh window function," in *IEEE Inter. Symp. on Circuits and Systems* (April 1974).
7. G.E. Loeb, "The functional replacement of the ear," *Sci. Am.* 252, 104-111 (1985).
8. T.Q. Nguyen, "The eigenfilter for the design of linear-phase filters with arbitrary magnitude response," in *ICASSP '91 Proceedings*, Vol 3, pp. 1981-1984 (May 1991).
9. T.W. Parks and J.H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Circuit Theory*, CT-19, 189-194 (1972).
10. G. Troullinos et al., "Theory and implementation of a splitband modem using the TMS32010," in P. Papamichalis (ed), *Digital Signal Processing Applications with the TMS320 Family - Volume 2*, Texas Instruments (1990), pp. 221-330.
11. R.S. Tyler, J.P. Preece, and M.W. Lowder, "The Iowa cochlear-implant test battery," Laboratory Report, University of Iowa at Iowa City, Department of Otolaryngology - Head and Neck Surgery, 1983.
12. B.S. Wilson et al., "Better speech recognition with cochlear implants," *Nature* 352, 236-238 (1991).
13. B.S. Wilson et al., "Speech processors for auditory prostheses," Tenth Quarterly Progress Report on NIH Contract NO1-DC-9-2401 (October 1991).
14. Working Group on Communication Aids for the Hearing-Impaired, "Speech-perception aids for hearing impaired people: current status and needed research," *J. Acoust. Soc. Amer.* 90, 637-683 (1991).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 12 January 1993		3. REPORT TYPE AND DATES COVERED Technical Report	
4. TITLE AND SUBTITLE Programmable Interactive System for Cochlear Implant Electrode Stimulation				5. FUNDING NUMBERS C — F19628-90-C-0002	
6. AUTHOR(S) Joseph Tierney, Marc A. Zissman, Donald K. Eddington, and William M. Rabinowitz					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Lincoln Laboratory, MIT P.O. Box 73 Lexington, MA 02173-9108				8. PERFORMING ORGANIZATION REPORT NUMBER TR-970	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) HQ Electronic Systems Division ESD/ENKL Hanscom AFB, MA 01730-5000				10. SPONSORING/MONITORING AGENCY REPORT NUMBER ESC-TR-92-13	
11. SUPPLEMENTARY NOTES None					
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.				12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The aim of this research, which was performed as a Lincoln Laboratory Innovative Research Program project, was to apply advanced digital speech and signal-processing techniques toward improving cochlear implant electrode stimulators. By providing a flexible stimulator whose function could be tuned depending on the subject's residual auditory nerves and the efficiency of the implant's coupling to those nerves, it was hypothesized that the subject's speech reception could be improved. The approach to providing these new and improved electrode stimulators included the design of a laboratory signal processor used for interactive testing of new algorithms with implant subjects. This Programmable Interactive System for Cochlear Implant Electrode Stimulation (PISCES) was designed, built, and tested at Lincoln Laboratory and then delivered to the Massachusetts Eye and Ear Infirmary (MEEI) Cochlear Implant Research Laboratory (CIRL). In collaboration with researchers at MEEI CIRL and MIT Research Laboratory of Electronics (RLE), new algorithms run on PISCES have resulted in substantial improvements in subject speech reception relative to that with their current implant stimulators. These results were obtained as a result of interactive algorithm adjustment at the clinic, which demonstrated the importance of a flexible signal processor.					
14. SUBJECT TERMS cochlear implant electrode stimulator speech processing continuous interleaved sampling hearing aid sensory deafness neural deafness auditory prosthesis				15. NUMBER OF PAGES 60	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Same as Report		

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102